

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/509,121
Source: Pg 110
Date Processed by STIC: 7/11/05

ENTERED



PCT

RAW SEQUENCE LISTING

DATE: 07/11/2005

PATENT APPLICATION: US/10/509,121

TIME: 15:06:55

Input Set : A:\2005-05-31 1254-0259PUS1.ST25.txt

Output Set: N:\CRF4\07112005\J509121.raw

3 <110> APPLICANT: Koichiro KAKU, et al.
 5 <120> TITLE OF INVENTION: GENE CODING FOR ACETOLACTATE SYNTHASE
 7 <130> FILE REFERENCE: 1254-0259PUS1
 9 <140> CURRENT APPLICATION NUMBER: US 10/509,121
 10 <141> CURRENT FILING DATE: 2004-09-28
 12 <150> PRIOR APPLICATION NUMBER: JP 2002-95721
 13 <151> PRIOR FILING DATE: 2002-03-29
 15 <160> NUMBER OF SEQ ID NOS: 39
 17 <170> SOFTWARE: PatentIn Ver. 2.0
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 2301
 21 <212> TYPE: DNA
 22 <213> ORGANISM: Oryza sativa var. kinmaze
 24 <220> FEATURE:
 25 <221> NAME/KEY: CDS
 26 <222> LOCATION: (48)..(1979)
 28 <400> SEQUENCE: 1

p.6

```

29 cccaaaccca gaaaccctcg ccgccgccgc cgccgccacc acccacc atg gct acg      56
30                                     Met Ala Thr
31                                     1
33 acc gcc gcg gcc gcg gcc gcc gcc ctg tcc gcc gcc gcg acg gcc aag      104
34 Thr Ala Ala Ala Ala Ala Ala Ala Leu Ser Ala Ala Ala Thr Ala Lys
35      5              10              15
37 acc ggc cgt aag aac cac cag cga cac cac gtc ctt ccc gct cga ggc      152
38 Thr Gly Arg Lys Asn His Gln Arg His His Val Leu Pro Ala Arg Gly
39 20              25              30              35
41 cgg gtg ggg gcg gcg gcg gtc agg tgc tcg gcg gtg tcc ccg gtc acc      200
42 Arg Val Gly Ala Ala Ala Val Arg Cys Ser Ala Val Ser Pro Val Thr
43      40              45              50
45 ccg ccg tcc ccg gcg ccg ccg gcc acg ccg ctc ccg ccg tgg ggg ccg      248
46 Pro Pro Ser Pro Ala Pro Pro Ala Thr Pro Leu Arg Pro Trp Gly Pro
47      55              60              65
49 gcc gag ccc cgc aag ggc gcg gac atc ctc gtg gag gcg ctg gag cgg      296
50 Ala Glu Pro Arg Lys Gly Ala Asp Ile Leu Val Glu Ala Leu Glu Arg
51      70              75              80
53 tgc ggc gtc agc gac gtg ttc gcc tac ccg ggc ggc gcg tcc atg gag      344
54 Cys Gly Val Ser Asp Val Phe Ala Tyr Pro Gly Gly Ala Ser Met Glu
55      85              90              95
57 atc cac cag gcg ctg acg cgc tcc ccg gtc atc acc aac cac ctc ttc      392
58 Ile His Gln Ala Leu Thr Arg Ser Pro Val Ile Thr Asn His Leu Phe
59 100              105              110              115
61 cgc cac gag cag ggc gag gcg ttc gcg gcg tcc ggg tac gcg cgc gcg      440
62 Arg His Glu Gln Gly Glu Ala Phe Ala Ala Ser Gly Tyr Ala Arg Ala
  
```

RAW SEQUENCE LISTING

DATE: 07/11/2005

PATENT APPLICATION: US/10/509,121

TIME: 15:06:55

Input Set : A:\2005-05-31 1254-0259PUS1.ST25.txt

Output Set: N:\CRF4\07112005\J509121.raw

63		120		125		130	
65	tcc ggc cgc gtc ggg gtc tgc gtc gcc acc tcc ggc ccc ggg gca acc						488
66	Ser Gly Arg Val Gly Val Cys Val Ala Thr Ser Gly Pro Gly Ala Thr						
67		135		140		145	
69	aac ctc gtg tcc gcg ctc gcc gac gcg ctg ctc gac tcc gtc ccg atg						536
70	Asn Leu Val Ser Ala Leu Ala Asp Ala Leu Leu Asp Ser Val Pro Met						
71		150		155		160	
73	gtc gcc atc acg ggc cag gtc cac agc cgc atg atc ggc acc gac gcc						584
74	Val Ala Ile Thr Gly Gln Val His Ser Arg Met Ile Gly Thr Asp Ala						
75		165		170		175	
77	ttc cag gag acg ccc ata gtc gag gtc acc cgc tcc atc acc aag cac						632
78	Phe Gln Glu Thr Pro Ile Val Glu Val Thr Arg Ser Ile Thr Lys His						
79	180		185		190		195
81	aat tac ctt gtc ctt gat gtg gag gac atc ccc cgc gtc ata cag gaa						680
82	Asn Tyr Leu Val Leu Asp Val Glu Asp Ile Pro Arg Val Ile Gln Glu						
83		200		205		210	
85	gcc ttc ttc ctc gcg tcc tcg ggc cgt cct ggc ccg gtg ctg gtc gac						728
86	Ala Phe Phe Leu Ala Ser Ser Gly Arg Pro Gly Pro Val Leu Val Asp						
87		215		220		225	
89	atc ccc aag gac atc cag cag cag atg gcc gtg ccg gtc tgg gac acc						776
90	Ile Pro Lys Asp Ile Gln Gln Gln Met Ala Val Pro Val Trp Asp Thr						
91		230		235		240	
93	tcg atg aat cta cca ggg tac atc gca cgc ctg ccc aag cca ccc gcg						824
94	Ser Met Asn Leu Pro Gly Tyr Ile Ala Arg Leu Pro Lys Pro Pro Ala						
95		245		250		255	
97	aca gaa ttg ctt gag cag gtc ttg cgt ctg gtt ggc gag tca cgg cgc						872
98	Thr Glu Leu Leu Glu Gln Val Leu Arg Leu Val Gly Glu Ser Arg Arg						
99	260		265		270		275
101	ccg att ctc tat gtc ggt ggt ggc tgc tct gca tct ggt gac gaa ttg						920
102	Pro Ile Leu Tyr Val Gly Gly Gly Cys Ser Ala Ser Gly Asp Glu Leu						
103		280		285		290	
105	cgc tgg ttt gtt gag ctg act ggt atc cca gtt aca acc act ctg atg						968
106	Arg Trp Phe Val Glu Leu Thr Gly Ile Pro Val Thr Thr Thr Leu Met						
107		295		300		305	
109	ggc ctc ggc aat ttc ccc agt gac gac ccg ttg tcc ctg cgc atg ctt						1016
110	Gly Leu Gly Asn Phe Pro Ser Asp Asp Pro Leu Ser Leu Arg Met Leu						
111		310		315		320	
113	ggg atg cat ggc acg gtg tac gca aat tat gcc gtg gat aag gct gac						1064
114	Gly Met His Gly Thr Val Tyr Ala Asn Tyr Ala Val Asp Lys Ala Asp						
115		325		330		335	
117	ctg ttg ctt gcg ttt ggt gtg cgg ttt gat gat cgt gtg aca ggg aaa						1112
118	Leu Leu Leu Ala Phe Gly Val Arg Phe Asp Asp Arg Val Thr Gly Lys						
119	340		345		350		355
121	att gag gct ttt gca agc agg gcc aag att gtg cac att gac att gat						1160
122	Ile Glu Ala Phe Ala Ser Arg Ala Lys Ile Val His Ile Asp Ile Asp						
123		360		365		370	
125	cca gca gag att gga aag aac aag caa cca cat gtg tca att tgc gca						1208
126	Pro Ala Glu Ile Gly Lys Asn Lys Gln Pro His Val Ser Ile Cys Ala						
127		375		380		385	

RAW SEQUENCE LISTING

DATE: 07/11/2005

PATENT APPLICATION: US/10/509,121

TIME: 15:06:55

Input Set : A:\2005-05-31 1254-0259PUS1.ST25.txt

Output Set: N:\CRF4\07112005\J509121.raw

```

129 gat gtt aag ctt gct tta cag ggc ttg aat gct ctg cta caa cag agc 1256
130 Asp Val Lys Leu Ala Leu Gln Gly Leu Asn Ala Leu Leu Gln Gln Ser
131      390      395      400
133 aca aca aag aca agt tct gat ttt agt gca tgg cac aat gag ttg gac 1304
134 Thr Thr Lys Thr Ser Ser Asp Phe Ser Ala Trp His Asn Glu Leu Asp
135      405      410      415
137 cag cag aag agg gag ttt cct ctg ggg tac aaa act ttt ggt gaa gag 1352
138 Gln Gln Lys Arg Glu Phe Pro Leu Gly Tyr Lys Thr Phe Gly Glu Glu
139 420      425      430      435
141 atc cca ccg caa tat gcc att cag gtg ctg gat gag ctg acg aaa ggt 1400
142 Ile Pro Pro Gln Tyr Ala Ile Gln Val Leu Asp Glu Leu Thr Lys Gly
143      440      445      450
145 gag gca atc atc gct act ggt gtt ggg cag cac cag atg tgg gcg gca 1448
146 Glu Ala Ile Ile Ala Thr Gly Val Gly Gln His Gln Met Trp Ala Ala
147      455      460      465
149 caa tat tac acc tac aag cgg cca cgg cag tgg ctg tct tcg gct ggt 1496
150 Gln Tyr Tyr Thr Tyr Lys Arg Pro Arg Gln Trp Leu Ser Ser Ala Gly
151      470      475      480
153 ctg ggc gca atg gga ttt ggg ctg cct gct gca gct ggt gct tct gtg 1544
154 Leu Gly Ala Met Gly Phe Gly Leu Pro Ala Ala Ala Gly Ala Ser Val
155      485      490      495
157 gct aac cca ggt gtc aca gtt gtt gat att gat ggg gat ggt agc ttc 1592
158 Ala Asn Pro Gly Val Thr Val Val Asp Ile Asp Gly Asp Gly Ser Phe
159 500      505      510      515
161 ctc atg aac att cag gag ctg gca ttg atc cgc att gag aac ctc cct 1640
162 Leu Met Asn Ile Gln Glu Leu Ala Leu Ile Arg Ile Glu Asn Leu Pro
163      520      525      530
165 gtg aag gtg atg gtg ttg aac aac caa cat ttg ggt atg gtg gtg caa 1688
166 Val Lys Val Met Val Leu Asn Asn Gln His Leu Gly Met Val Val Gln
167      535      540      545
169 tgg gag gat agg ttt tac aag gcg aat agg gcg cat aca tac ttg ggc 1736
170 Trp Glu Asp Arg Phe Tyr Lys Ala Asn Arg Ala His Thr Tyr Leu Gly
171      550      555      560
173 aac ccg gaa tgt gag agc gag ata tat cca gat ttt gtg act att gct 1784
174 Asn Pro Glu Cys Glu Ser Glu Ile Tyr Pro Asp Phe Val Thr Ile Ala
175      565      570      575
177 aag ggg ttc aat att cct gca gtc cgt gta aca aag aag agt gaa gtc 1832
178 Lys Gly Phe Asn Ile Pro Ala Val Arg Val Thr Lys Lys Ser Glu Val
179 580      585      590      595
181 cgt gcc gcc atc aag aag atg ctc gag act cca ggg cca tac ttg ttg 1880
182 Arg Ala Ala Ile Lys Lys Met Leu Glu Thr Pro Gly Pro Tyr Leu Leu
183      600      605      610
185 gat atc atc gtc ccg cac cag gag cat gtg ctg cct atg atc cca agt 1928
186 Asp Ile Ile Val Pro His Gln Glu His Val Leu Pro Met Ile Pro Ser
187      615      620      625
189 ggg ggc gca ttc aag gac atg atc ctg gat ggt gat ggc agg act gtg 1976
190 Gly Gly Ala Phe Lys Asp Met Ile Leu Asp Gly Asp Gly Arg Thr Val
191      630      635      640
193 tat taatctataa tctgtatggt ggcaaagcac cagcccgcc tatgtttgac 2029

```

RAW SEQUENCE LISTING

DATE: 07/11/2005

PATENT APPLICATION: US/10/509,121

TIME: 15:06:55

Input Set : A:\2005-05-31 1254-0259PUS1.ST25.txt

Output Set: N:\CRF4\07112005\J509121.raw

```

194 Tyr
196 ctgaatgacc cataaagagt ggtatgccta tgatgtttgt atgtgctcta tcaataacta 2089
198 aggtgtcaac tatgaaccat atgctcttct gttttacttg tttgatgtgc ttggcatggg 2149
200 aatcctaatt agcttcctgc tgtctagggt tgtagtgtgt tgttttctgt aggcatatgc 2209
202 atcacaagat atcatgtaag tttcttgtcc tacatatcaa taataagaga ataaagtact 2269
204 tctatgcaaa aaaaaaaaaa aaaaaaaaaa aa 2301
207 <210> SEQ ID NO: 2
208 <211> LENGTH: 644
209 <212> TYPE: PRT
210 <213> ORGANISM: Oryza sativa var. kinmaze
212 <400> SEQUENCE: 2
213 Met Ala Thr Thr Ala Ala Ala Ala Ala Ala Leu Ser Ala Ala Ala
214 1 5 10 15
216 Thr Ala Lys Thr Gly Arg Lys Asn His Gln Arg His His Val Leu Pro
217 20 25 30
219 Ala Arg Gly Arg Val Gly Ala Ala Ala Val Arg Cys Ser Ala Val Ser
220 35 40 45
222 Pro Val Thr Pro Pro Ser Pro Ala Pro Pro Ala Thr Pro Leu Arg Pro
223 50 55 60
225 Trp Gly Pro Ala Glu Pro Arg Lys Gly Ala Asp Ile Leu Val Glu Ala
226 65 70 75 80
228 Leu Glu Arg Cys Gly Val Ser Asp Val Phe Ala Tyr Pro Gly Gly Ala
229 85 90 95
231 Ser Met Glu Ile His Gln Ala Leu Thr Arg Ser Pro Val Ile Thr Asn
232 100 105 110
234 His Leu Phe Arg His Glu Gln Gly Glu Ala Phe Ala Ala Ser Gly Tyr
235 115 120 125
237 Ala Arg Ala Ser Gly Arg Val Gly Val Cys Val Ala Thr Ser Gly Pro
238 130 135 140
240 Gly Ala Thr Asn Leu Val Ser Ala Leu Ala Asp Ala Leu Leu Asp Ser
241 145 150 155 160
243 Val Pro Met Val Ala Ile Thr Gly Gln Val His Ser Arg Met Ile Gly
244 165 170 175
246 Thr Asp Ala Phe Gln Glu Thr Pro Ile Val Glu Val Thr Arg Ser Ile
247 180 185 190
249 Thr Lys His Asn Tyr Leu Val Leu Asp Val Glu Asp Ile Pro Arg Val
250 195 200 205
252 Ile Gln Glu Ala Phe Phe Leu Ala Ser Ser Gly Arg Pro Gly Pro Val
253 210 215 220
255 Leu Val Asp Ile Pro Lys Asp Ile Gln Gln Gln Met Ala Val Pro Val
256 225 230 235 240
258 Trp Asp Thr Ser Met Asn Leu Pro Gly Tyr Ile Ala Arg Leu Pro Lys
259 245 250 255
261 Pro Pro Ala Thr Glu Leu Leu Glu Gln Val Leu Arg Leu Val Gly Glu
262 260 265 270
264 Ser Arg Arg Pro Ile Leu Tyr Val Gly Gly Gly Cys Ser Ala Ser Gly
265 275 280 285
267 Asp Glu Leu Arg Trp Phe Val Glu Leu Thr Gly Ile Pro Val Thr Thr
268 290 295 300

```

RAW SEQUENCE LISTING

DATE: 07/11/2005

PATENT APPLICATION: US/10/509,121

TIME: 15:06:55

Input Set : A:\2005-05-31 1254-0259PUS1.ST25.txt

Output Set: N:\CRF4\07112005\J509121.raw

```

270 Thr Leu Met Gly Leu Gly Asn Phe Pro Ser Asp Asp Pro Leu Ser Leu
271 305                               310                   315           320
273 Arg Met Leu Gly Met His Gly Thr Val Tyr Ala Asn Tyr Ala Val Asp
274                               325                   330           335
276 Lys Ala Asp Leu Leu Leu Ala Phe Gly Val Arg Phe Asp Asp Arg Val
277                               340                   345           350
279 Thr Gly Lys Ile Glu Ala Phe Ala Ser Arg Ala Lys Ile Val His Ile
280                               355                   360           365
282 Asp Ile Asp Pro Ala Glu Ile Gly Lys Asn Lys Gln Pro His Val Ser
283                               370                   375           380
285 Ile Cys Ala Asp Val Lys Leu Ala Leu Gln Gly Leu Asn Ala Leu Leu
286 385                               390                   395           400
288 Gln Gln Ser Thr Thr Lys Thr Ser Ser Asp Phe Ser Ala Trp His Asn
289                               405                   410           415
291 Glu Leu Asp Gln Gln Lys Arg Glu Phe Pro Leu Gly Tyr Lys Thr Phe
292                               420                   425           430
294 Gly Glu Glu Ile Pro Pro Gln Tyr Ala Ile Gln Val Leu Asp Glu Leu
295                               435                   440                   445
297 Thr Lys Gly Glu Ala Ile Ile Ala Thr Gly Val Gly Gln His Gln Met
298                               450                   455                   460
300 Trp Ala Ala Gln Tyr Tyr Thr Tyr Lys Arg Pro Arg Gln Trp Leu Ser
301 465                               470                   475           480
303 Ser Ala Gly Leu Gly Ala Met Gly Phe Gly Leu Pro Ala Ala Ala Gly
304                               485                   490           495
306 Ala Ser Val Ala Asn Pro Gly Val Thr Val Val Asp Ile Asp Gly Asp
307                               500                   505           510
309 Gly Ser Phe Leu Met Asn Ile Gln Glu Leu Ala Leu Ile Arg Ile Glu
310                               515                   520                   525
312 Asn Leu Pro Val Lys Val Met Val Leu Asn Asn Gln His Leu Gly Met
313                               530                   535           540
315 Val Val Gln Trp Glu Asp Arg Phe Tyr Lys Ala Asn Arg Ala His Thr
316 545                               550                   555           560
318 Tyr Leu Gly Asn Pro Glu Cys Glu Ser Glu Ile Tyr Pro Asp Phe Val
319                               565                   570           575
321 Thr Ile Ala Lys Gly Phe Asn Ile Pro Ala Val Arg Val Thr Lys Lys
322                               580                   585           590
324 Ser Glu Val Arg Ala Ala Ile Lys Lys Met Leu Glu Thr Pro Gly Pro
325                               595                   600           605
327 Tyr Leu Leu Asp Ile Ile Val Pro His Gln Glu His Val Leu Pro Met
328                               610                   615           620
330 Ile Pro Ser Gly Gly Ala Phe Lys Asp Met Ile Leu Asp Gly Asp Gly
331 625                               630                   635           640
333 Arg Thr Val Tyr
337 <210> SEQ ID NO: 3
338 <211> LENGTH: 2300
339 <212> TYPE: DNA
340 <213> ORGANISM: Oryza sativa var. kinmaze
342 <220> FEATURE:
343 <221> NAME/KEY: CDS

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 07/11/2005
PATENT APPLICATION: US/10/509,121 TIME: 15:06:56

Input Set : A:\2005-05-31 1254-0259PUS1.ST25.txt
Output Set: N:\CRF4\07112005\J509121.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:29; N Pos. 15

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/509,121

DATE: 07/11/2005

TIME: 15:06:56

Input Set : A:\2005-05-31 1254-0259PUS1.ST25.txt

Output Set: N:\CRF4\07112005\J509121.raw

L:1545 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29 after pos.:0